

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:08 ; Search time 210.42 Seconds

(without alignments)
11.213 Million cell updates/sec

Title: US-09-331-631a-3_COPY_117_185

Perfect score: 384
Sequence: 1 NKGRDPGQGYEQCGECGRH.....EEQREDEKYEERKKEEDN 69

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36: *
1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT: *
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT: *
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT: *
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT: *
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT: *
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT: *
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT: *
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT: *
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT: *
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT: *
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT: *
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT: *
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT: *
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT: *
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT: *
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT: *
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT: *
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT: *
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT: *
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT: *
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	384	100.0	666	19 W62829	Macadamia integrifl
2	367	95.6	666	19 W62828	Macadamia integrifl
3	359	93.5	625	19 W62830	Macadamia integrifl
4	175	45.6	525	19 W62831	Theobroma cacao an
5	175	45.6	566	13 R20181	Sequence encoded b
6	149	38.8	590	19 W62832	Gossypium hirsutum
7	113	29.4	1898	20 Y30795	A human trichohyal
8	111	28.9	1162	21 Y58500	HHV8 ORF 73 protei
9	103.5	27.0	409	20 W90342	G. max truncated S
10	103.5	27.0	489	20 W90341	G. max SBP2 protei
11	102.5	26.7	611	20 Y29039	T. gondii Immunoge
12	99	25.8	1135	21 Y68784	Amino acid sequenc

13	99	25.8	1233	20 Y55954	Mouse STE20-relate
14	99	25.8	1239	20 Y55931	Human ZC1 protein.
15	98.5	25.7	1297	20 Y55932	Human ZC2 protein.
16	98.5	25.7	1360	21 Y85263	Human protein kina
17	98	25.5	482	20 Y07067	Renal cancer assoc
18	98	25.5	562	16 R70491	Leucocytozan assoc
19	97	25.3	2023	21 Y54320	Human acid sequenc
20	96	25.0	444	20 W90339	G. max truncated S
21	96	25.0	524	20 W90339	G. max SBP1 protei
22	94.5	24.6	1326	20 Y55933	Human ZC3 protein.
23	92	24.0	2074	21 Y54319	Amino acid sequenc
24	91	23.7	360	17 W03627	Human foliicic scl
25	91	23.7	412	17 W03626	Human thylotrophi
26	91	23.7	593	19 W62835	Zea mays antimicro
27	90.5	23.6	740	13 R27530	Plasmodium falci
28	90.5	23.6	740	16 R68838	Plasmodium falci
29	90.5	23.6	1299	21 Y58633	Protein regulating
30	88.5	23.0	303	15 R60054	Dirotifilaria imm
31	87.5	22.8	346	20 Y20115	B. burgdorferi ant
32	87.5	22.8	373	20 Y20114	Human p160 polypep
33	87.5	22.8	505	18 W31186	Human p160 polypep
34	87.5	22.8	1135	18 W31185	Human regulatory m
35	87	22.7	432	20 W93954	Human regulatory m
36	86	22.4	386	11 R06849	Protein Arp 4. R
37	86	22.4	386	11 R04828	Arp 4. Streptococ
38	85	22.4	910	20 Y22191	Mouse brain CNG-1
39	85	22.1	1178	18 W30763	Mannose-1-phosphat
40	84.5	22.0	186	18 W26536	Trypanosoma cruzi
41	84.5	22.0	186	20 Y23298	Trypanosoma cruzi
42	84.5	22.0	301	8 P70867	Sequence of acidic
43	84.5	22.0	326	20 Y20119	B. burgdorferi ant
44	84.5	22.0	359	20 Y20118	B. burgdorferi ant
45	84.5	22.0	605	20 Y40999	Soybean beta-congl

ALIGNMENTS

RESULT 1	
W62829	
ID	W62829 standard; Protein; 666 AA.
AC	W62829;
XX	
DT	27-OCT-1998 (first entry)
DE	Macadamia integrifolia antimicrobial protein.
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	/note="signal peptide"
FT	29..666
FT	/note="mature protein"
XX	
PN	W09827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97WO-AU00874.
XX	
PR	20-DEC-1996; 96AU-0004275.
XX	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX	WPI: 1998-377279/32.
DR	N-PSDB: V42311.
XX	

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 39-41; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
SQ Sequence 666 AA;

Query Match 100.0%; Score 384; DB 19; Length 666;
Best Local Similarity 100.0%; Pred. No. 5.1e-31;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NRORDPQOOYEEOCERQHETPRHMOTCOQCERRYEKERKROOKRYEEOQREDEEKY 60
DB 117 nrgtdpqyqeqcqrtrreptmqlcqcqrcerryekrkqkryeeqgredexy 176
OY 61 EERMKEDN 69
DB 177 eermkeedn 185

RESULT 2

W62828 standard; Protein; 666 AA.

W62828;

27-OCT-1998 (first entry)

Macadamia integrifolia antimicrobial protein.

antimicrobial protein; infestation; control.

Macadamia integrifolia.

Key Location/Qualifiers

FT Peptide 1..28 /note="signal peptide"

FT Protein 29..666 /note="mature protein"

W09827805-A1.

02-JUL-1998.

22-DEC-1997; 97WO-AU00874.

20-DEC-1996; 96AU-0004275.

(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

WPI; 1998-377279/32.

N-PSDB; VA2310.

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 34-36; 96pp; English.

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.

SQ Sequence 666 AA;

Query Match 95.6%; Score 367; DB 19; Length 666;

Best Local Similarity 95.7%; Pred. No. 2.6e-29;
Matches 66; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NRORDPQOOYEEOCERQHETPRHMOTCOQCERRYEKERKROOKRYEEOQREDEEKY 60
DB 117 nrgtdpqyqeqcqrtrreptmqlcqcqrcerryekrkqkryeeqgredexy 176

OY 61 EERMKEDN 69
DB 177 eermkeedn 185

RESULT 3

W62830 standard; Protein; 625 AA.

W62830;

27-OCT-1998 (first entry)

Macadamia integrifolia antimicrobial protein.

antimicrobial protein; infestation; control.

Macadamia integrifolia.

Key Location/Qualifiers

FT Peptide 1..28 /note="signal peptide"

FT Protein 29..666 /note="mature protein"

W09827805-A1.

02-JUL-1998.

22-DEC-1997; 97WO-AU00874.

20-DEC-1996; 96AU-0004275.

(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

WPI; 1998-377279/32.

N-PSDB; VA2316.

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 43-45; 96pp; English.

CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.

SQ Sequence 625 AA;

Query Match 93.5%; Score 359; DB 19; Length 625;
Best Local Similarity 94.2%; Pred. No. 1.5e-28;
Matches 65; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 NRORDPQOOYEEOCERQHETPRHMOTCOQCERRYEKERKROOKRYEEOQREDEEKY 60
DB 76 nrgtdpqyqeqcqrtrreptmqlcqcqrcerryekrkqkryeeqgredexy 135

OY 61 EERMKEDN 69
DB 136 eermkeedn 144

RESULT 4

```

W62831
ID W62831 standard; Protein; 525 AA.
XX
AC W62831:
XX
DT 27-OCT-1998 (first entry)
XX
DE Theobroma cacao antimicrobial protein.
XX
KW antimicrobial protein; infestation; control.
XX
OS Theobroma cacao.
XX
PN MO9827805-A1.
XX
PD 02-JUL-1998.
XX
PE 22-DEC-1997; 97MO-AU00874.
XX
PR 20-DEC-1996; 96AU-0004275.
XX
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR WPI: 1998-377279/32.
XX
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX
PS useful for controlling microbial infestations of plants or mammals
XX
CC Claim 1; Page 47-49; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
XX
CC be used to control microbial infestations in plants and mammalian
XX
CC animals.
XX
SQ Sequence 525 AA;

Query Match 45.6%; Score 175; DB 19; Length 525;
Best Local Similarity 32.4%; Pred. No. 3,4e-10;
Matches 33; Conservative 19; Mismatches 14; Indels 36; Gaps 2;

QY 3 QRDPOQYEQCOERCQHEHTEPRHMQTOQRCERRYEKKRKO----- 46
   :||:||||||| || : ||||| :||:|:|
Dy 35 erdpqygeqgcrcesateeregeqgcrcereykedqrgqgeelqrgyqgcgrcge 94
   :||:||||||| || : ||||| :||:|:|
QY 47 -----KRYEQQRDEDEKY----ERRKKEED 68
   :||:||||| :| : ||||
Dy 95 qqqgqrgqgcqrcqkwegykqergerhenyinhkknrseeee 136

RESULT 5
R20181
ID R20181 standard; Protein; 566 AA.
XX
AC R20181:
XX
DT 16-APR-1992 (first entry)
XX
DE Sequence encoded by 67 kD T. cacao protein cDNA.
XX
KW Cocoa; flavour; vicillin; seed storage protein.
XX
OS Theobroma cacao.
XX
PN WO9119801-A.
XX
PD 26-DEC-1991.
XX
PE 07-JUN-1991; 91WO-GB00914.
XX
PR 11-JUN-1990; 90GB-0013016.

```

```

XX RA      (MASC ) MARKS UK LTD.
XX PI      Spencer ME, Hodge R, Deakin EA, Ashton S;
XX DR      WPI: 1992-024418/03.
XX N-PSDB: Q20377.
XX PT      Recombinant cocoa proteins - are responsible for flavour in cocoa
XX expression vectors
XX PS
XX Claim 4; Fig 2; 59pp; English.
XX
XX CC      The inventors claim a 67 kD and 31 kD T. cacao protein, and
XX CC      fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
XX CC      derived from the 67 kD precursor. T. cacao protein cDNA was
XX CC      detected in a cDNA library prepared from immature cocoa beans RNA
XX CC      using a probe based on the AA sequence of a CNBR peptide common to
XX CC      the 47 kD and 31 kD polypeptides. Homology searches revealed close
XX CC      homologues between the 67 kD polypeptide and the vicilins, which are
XX CC      seed storage proteins.
XX SO      Sequence     566 AA;
XX
XX Query Match          45.6%; Score 175; DB 13; Length 566;
XX Best Local Similarity 32.4%; Pred.No. 3.7e-10;
XX Matches   33; Conservative 19; Mismatches 14; Indels   36; Gaps
XX
XX QY      3 QRDPQOQYECCGKERQRHETPRHMQTQQCGRRRYEKKRKO----- 46
XX           :|||:||||| || | : ||||| |:::||:
XX Db       35 erdprqyqcqricesateeteegcgcrcrerykqqrgeelrqygqcgqrcqe 94
XX QY      47 -----KRYEQQRDEEKY----ERRKKEED 68
XX           ::|||::| :| :| :| :| :|
XX Db       95 qgggqrgqqgcqirkewgykeqergghenyhmhkknreeee 136
XX
XX RESULT    6
XX ID        W62832 standard; protein; 590 AA.
XX AC        W62832;
XX DT        27-OCT-1998 (first entry)
XX DE        Gossypium hirsutum antimicrobial protein.
XX KW        antimicrobial protein; infestation; control.
XX OS        Gossypium hirsutum.
XX PN        W09827805-A1.
XX PD        02-JUL-1998.
XX PF        22-DEC-1997; 97MO-AU00874.
XX PR        20-DEC-1996; 96AU-0004275.
XX PA        (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX PI        Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX DR        WPI: 1998-377279/32.
XX PT        Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX useful for controlling microbial infestations of plants or mammals
XX PS        Claim 1; Page 49-51; 96pp; English.
XX TC        The sequence is that of an antimicrobial protein which can

```

CC be used to control microbial infestations in plants and mammalian
 CC animals.
 XX
 SQ Sequence 590 AA;

Query Match	38.8%	Score 149	DB 19	Length 590
Best Local Similarity	33.0%	Pred. No. 1.6e-07		
Matches 31	Conservative 19	Mismatches 16	Indels 28	Gaps 3

```

QY      1  NRQDPPQDQTECCQCEQCRHNETEPRHMQTCQDQKCEKRYEKEKKRQDQKRYEQD----- 53
      :  |||::|||::|  :  :  :  ||||  :||::|  :  :  :  :
DB      78  hrpdpqrtyeegqecr--qgeerqpqcgqrclkrfegedqgsqrfqecqchbqge 135

```

```
QY 54 -----REDEEKY-----EERMKED 68
      || |||
Db 136 grpekkgcvrecrekyqenpwrgerereeeeee 169
```

RESULT	7
Y30795	
ID	Y30795 standard; Protein; 1898 AA.
vv	

DT 25-NOV-1999 (first entry)
XX
DE A human trichohyalin (TRHY) protein.

KW Human; trichohyalin; TRHY; protein; tissue structure; wound healing;
KW terminally differentiating epidermal tissue; proteinaceous gel;
KW breast implant.

OS	Homo sapiens.
XX	
PN	US5958752-A.

PD	28-SEP-1999.
XX	
PF	14-FEB-1997; 97US-0800644.

PR 30-APR-1993; 93US-0056200.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

Kim I, Chung S, Park S, Steinert PM, Lee S, WPI; 1999-561041/47.

XX	Human trichohyalin useful for forming a proteinaceous gel that promotes
PT	wound healing -
PT	
DN	IN F508, 042301.

aa Disclosure; Fig 3A-W; 126pp; English.
ps
ps
xx The present sequence represents a human
cc

CC The present sequence represents a human trichohyalin (TRH) protein.
CC The protein is found in terminally differentiating epidermal tissue,
CC and is involved in forming the structural architecture of such
CC tissue. The trichohyalin protein is useful for forming a
CC proteanaceous gel which may then be used for healing wounds, or in
CC breast implants.

Sequence	1898	AA;
SQ		

Query Match	29.48;	Score 113;	DB 20;	Length 1898;
Best Local Similarity	33.38;	Pred. 0.0022;		
Matches 25; Conservative	22;	Mismatches 20;	Indels 8;	Gaps 2

```

OY 2 R0NDP000YEQCQERKORHETPRHMQTCQQRERRYEK-----EKRR0QKRYEEQ-Q 53
    |::: ::::: |::| | |::: |::| |::: |::| |::: |::| |:::
Db 468 reqeerrtdw|kreeeterherbrkq|krdgeerterw|kleeeerreqerreq| 527

```

```
QY      54 REDEEKYEERMKEED 68  
         ||||:|:|:|:  
Db      528 regeerreqrlkrqe 542
```

RESULT	8
Y58500	
ID	Y58500 standard; Protein; 1162 AA.

AC	Y58500;
XX	
DT	10-APR-2000 (first entry)

DE	HHV8 ORF 73 protein, SEQ ID NO:21.
XX	
KW	HHV8; detection; diagnosis; Kaposi

KW antigen.
XX
OS Human herpesvirus type 8.

	Key	Location/Qualifiers
FH	Misc-difference	96
FT		
FT		/label= unknown
FT		

PN	WO9961909-A2.
XX	
PD	02-DEC-1999.
UN	

PF	26-MAY-1999;	99WO-US11407.
XX		
PR	26-MAY-1998;	98US-0086695.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES
XX
PI Pau CP;

DR	WPI; 2000-097142/08.
XX	
PT	New methods and comp

PS Claim 2; Page 59-62; 68pp; English.

Sequence Y58480-Y58532 represent immunogenic oligopeptides derived from human herpes virus type 8 (HHV8, a gammaherpesvirus). HHV8 plays an important role in the pathogenesis of AIDS-related Kaposi's sarcoma. The invention relates to a novel method of detecting the presence of human herpesvirus 8 in a biological sample using peptides representative of dominant antigenic regions of HHV8. The method comprises contacting one or more isolated, immunogenic HHV8 peptides with an antibody-containing biological sample, and detecting the formation of a complex between the peptide and the antibody. The presence of a peptide-antibody complex indicates the presence of human herpesvirus 8. The detection of HHV8 infection can be used to diagnose AIDS-associated Kaposi's sarcoma. The HHV8-specific antibodies are useful therapeutically when for the passive immunisation of a human against HHV8 infection, thereby reducing HHV8 related disease. The detection assays are highly specific, sensitive and accurate. Early detection and treatment of Kaposi's sarcoma could diminish the severity of symptoms related to AIDS and the sensitive techniques could reduce erroneous characterisations of skin disorders. Previous assays for HHV8 antibodies such as immunofluorescence assays, immunoblots and enzyme immunoassays lack the sensitivity and accuracy needed for reliable diagnosis of Kaposi's sarcoma. Further advantages of the assays are that reproducible results are obtained and the method is suitable for rapid throughput and screening of samples economically.

Sequence 1162 AA;

Query Match	28.9%	Score 111	DB 21	Length 1162
Best Local	31.3%	Pred No. 0.0021		
Matches 21	Conservative 26	Mismatches 20	Indels 0	Gaps 0

2 RQRDPQQYECCQRCQRHETEPKHMOTCQQRCEKRYEKEKRYEEOQREDEKYE 61

KM Immunogenic protein; Toxoplasma gondii protein; oocyst shedding; cat;
 KW T. gondii infection; enteric apicomplexa oocyst; Cryptosporidium oocyst;
 XX Toxoplasma oocyst.
 OS Toxoplasma gondii.
 XX
 PM W09932633-A1.
 PD
 XX 01-JUL-1999.
 XX
 PF 18-DEC-1998; 98WO-US27137.
 XX
 PR 19-DEC-1997; 97US-0994825.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Lutz SB, Milhausen MJ, Ng RK;
 XX
 DR WPI; 1999-418930/35.
 DR N-PSDB; X91242.
 XX
 PT New Isolated Toxoplasma gondii nucleic acids used, e.g. to treat
 PT infection caused by this microorganism
 XX
 Claim 29; Page 227-229; 381pp; English.
 PS
 XX
 CC The invention provides isolated Toxoplasma gondii nucleic acids that
 CC encode immunogenic polypeptides. The T. gondii nucleic acid molecules,
 CC immunogenic proteins and antibodies to the proteins can be used to
 CC inhibit T. gondii oocyst shedding in a cat due to infection with
 CC T. gondii. They can be used for preventing T. gondii infection and for
 CC preventing the spread of T. gondii infection. They can also be used for
 CC detecting T. gondii infection. The detection method can be used to detect
 CC parasite cysts or oocysts in feces, e.g. from enteric apicomplexa oocysts
 CC such as Cryptosporidium oocysts and Toxoplasma oocysts.
 XX
 Sequence 611 AA;

Query Match	26.7%	Score 102.5;	DB 20;	Length 611;
Best Local Similarity	30.0%	Pred. No. 0.0073;		
Matches 21; Conservative	22;	Mismatches 24;	Indels 3;	Gaps 1;

```
QY      2  RORP000OYBCCSORC0HNEPFRPHMTCOORCRRRYEKEXKROOKRYEED---QREDEE 58
      | : | : : : | | | : : : | : | : : : | : | : | : |
Db      344 reedererrrveeekarqreereerrrveeekarqreeseerrrveeekarqreese 403

QY      59 KYEERMKEED 68
      : | | | |
Db      404 errerrrvee 413
```

RESULT	12
ID	Y68784
	Y68784 standard; Protein: 1135 AA.
XX	
AC	Y68784;
XX	
DT	16-MAY-2000 (first entry)
XX	
DE	Amino acid sequence of a human phosphorylation effector PHSP-16
XX	
KW	Human: phosphorylation effector; PHSF; proliferative disorder;
KM	immune disorder; neuronal disorder.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Modified-site
FT	/note= "potential phosphorylation site"
FT	Modified-site
FT	17
FT	/note= "potential phosphorylation site"
Region	31..54

FT	Modified-site	/note-33	"protein kinase signature sequence"
FT	Modified-site	/note-59	"potential glycosylation site"
FT	Modified-site	/note-59	"potential phosphorylation site"
FT	Modified-site	/note-77	"potential phosphorylation site"
FT	Modified-site	/note-112	"potential phosphorylation site"
FT	Modified-site	/note-124	"potential phosphorylation site"
FT	Region	/note-129..182	"potential phosphorylation site"
FT	Region	/note-149..161	"protein kinase signature sequence"
FT	Modified-site	/note-187	"protein kinase signature sequence"
FT	Active-site	/note-190..200	"potential phosphorylation site"
FT	Active-site	/note-211..236	"tyrosine kinase catalytic site"
FT	Modified-site	/note-222	"tyrosine kinase catalytic site"
FT	Modified-site	/note-255	"potential phosphorylation site"
FT	Modified-site	/note-259	"potential phosphorylation site"
FT	Modified-site	/note-264	"potential phosphorylation site"
FT	Modified-site	/note-309	"potential phosphorylation site"
FT	Modified-site	/note-319	"potential phosphorylation site"
FT	Modified-site	/note-321	"potential phosphorylation site"
FT	Modified-site	/note-323	"potential phosphorylation site"
FT	Modified-site	/note-324	"potential phosphorylation site"
FT	Modified-site	/note-326	"potential phosphorylation site"
FT	Modified-site	/note-351	"potential phosphorylation site"
FT	Modified-site	/note-467	"potential phosphorylation site"
FT	Modified-site	/note-543	"potential phosphorylation site"
FT	Modified-site	/note-550	"potential phosphorylation site"
FT	Modified-site	/note-554	"potential phosphorylation site"
FT	Modified-site	/note-570	"potential phosphorylation site"
FT	Modified-site	/note-572	"potential glycosylation site"
FT	Modified-site	/note-624	"potential phosphorylation site"
FT	Modified-site	/note-625	"potential phosphorylation site"
FT	Modified-site	/note-632	"potential phosphorylation site"
FT	Modified-site	/note-661	"potential phosphorylation site"
FT	Modified-site	/note-662	"potential phosphorylation site"
FT	Modified-site	/note-688	"potential phosphorylation site"
FT	Modified-site	/note-669	"potential phosphorylation site"
FT	Modified-site	/note-706	"potential phosphorylation site"
FT	Modified-site	/note-	"potential phosphorylation site"

FT	Modified-site	718	/note=	"potential glycosylation site"
FT	Modified-site	720	/note=	"potential phosphorylation site"
FT	Modified-site	726	/note=	"potential phosphorylation site"
FT	Modified-site	811	/note=	"potential phosphorylation site"
FT	Modified-site	815	/note=	"potential phosphorylation site"
FT	Domain	836..1115	/note=	"potential phosphorylation site"
FT	Modified-site	898	/note=	"NIK1-like kinase domain"
FT	Modified-site	931	/note=	"potential phosphorylation site"
FT	Modified-site	958	/note=	"potential phosphorylation site"
FT	Modified-site	978	/note=	"potential phosphorylation site"
FT	Modified-site	999	/note=	"potential phosphorylation site"
FT	Modified-site	1012	/note=	"potential phosphorylation site"
FT	Modified-site	1067	/note=	"potential phosphorylation site"
FT	Modified-site	1113	/note=	"potential glycosylation site"
FT	Modified-site		/note=	"potential phosphorylation site"
XX				
PM	W0200006728-A2.			
PD	10-FEB-2000.			
XX				
PE	28-JUL-1999;	99WO-US17132.		
XX				
XX	28-JUL-1998;	98US-0123494.		
PR	14-SEP-1998;	98US-0152814.		
PR	14-OCT-1998;	98US-0173482.		
PR	03-NOV-1998;	98US-0106889.		
PR	19-NOV-1998;	98US-0109093.		
PR	22-DEC-1998;	98US-0113796.		
PR	12-JAN-1999;	98US-0173482.		
PR	12-JAN-1999;	99US-0229005.		
XX				
PA	(INCY-)	INCYTE PHARM INC.		
XX				
PI	Hillman JL, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;			
PI	Patterson C, Bandman O, Au-Young J, Gorgone GA, Yue H, Azimzal Y;			
PI	Reddy R, Lu DAM, Shih LT;			
XX				
DR	WPI: 2000-183125/16.			
DR	N-PSDB: Z46153.			
XX				
XX	New human phosphorylation effectors useful for the diagnosis, treatment			
PT	and prevention of proliferative, immune and neuronal disorders -			
XX				
PS	Claim 1; Page 98-100; 142pp; English.			
XX				
CC	Y68769-95 and Y68797-99 represent human phosphorylation effectors (PHSP)			
CC	designated PHSP1-PHSP1 (the protein sequence for PHSP28 is not given			
CC	in the specification). The sequences were isolated from cDNA libraries			
CC	prepared from various human tissues. The PHSP proteins are useful for			
CC	the diagnosis, treatment and prevention of proliferative disorders,			
CC	immune disorders and neuronal disorders. The PHSP proteins form			
CC	pharmaceutical compositions which useful for treating or preventing			
CC	disorders associated with decreased PHSP expression/activity. PHSP			
CC	antagonists are useful for treating or preventing disorders associated			
XX	with increased PHSP expression/activity.			
XX				
SO	Sequence 1135 AA;			

	Best Local Similarity	36.1%;	Pred. No. 0.032;	
	Matches	30; Conservative	15; Mismatches	20; Indels
Oy	2 RORPQQOYEECC--EKCQNHTEPRHMYCQQRRCRRYEKKERK-----QKRVEEQCR- 54	: :	:	: :
Dd	395 rqrkrlqekqrllrrleeqqrrearrqrqrgr--rfegeekrlleelerrkkeerrr 452	:	:	:
Oy	55 -----EDEEKVEERKEED 68	: :	:	:
Dd	453 raeeekrrrvreegyirrrgleec 475	:	:	:
	RESULT 13			
ID	y55954 standard: Protein; 1233 AA.			
XX	AC y55954:			
XX	DT 18-FEB-2000 (first entry)			
DE	Mouse STE20-related protein kinase NIK_m.			
XX	Antirheumatic; antithyritic; antiinflammatory; antiallergic; osteopathic;			
XX	antiporiatic; antiatherosclerotic; antiasthmatic; immunosuppressive;			
KW	neuroprotective; cardiact; cerebroprotective; cytostatic; antidiabetic;			
KV	vulneray: STE20; protein kinase: STLK2; STLK3; STLK5; STLK6; STLK7			
KM	ZC1, ZC2, ZC3, ZC4, KHS2, SNUJ1, SNUJ3, GEX2, PAK4, PAK5; antagonist;			
KM	antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;			
KM	inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;			
KM	rhlntils; autoimmunity; organ transplantation; multiple sclerosis;			
KM	amyocardial infarction; cardiovascular disease; stroke; renal failure;			
KM	oxidative stress-related neurodegenerative disorder; Parkinson's disease			
KM	amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;			
KM	ischemic disorder; inflammation; diabetes mellitus; fibrosis; mictosis;			
KM	mesangial disorder; growth regulation; wound healing; T cell activation;			
XX	immunosuppressant.			
XX	Mus sp.			
OS	WO9953036-A2.			
PN	21-OCT-1999.			
PD	13-APR-1999: 99MO-US08150.			
PF	14-APR-1998: 98US-0081784.			
XX	(SUGEN-) SUGEN INC.			
XX	Plowman G, Martinez R, Whyte D;			
XX	WPt: 1999-611301/52.			
DR	Novel kinase-related polypeptides used for the diagnosis and treatment			
PT	of kinase-related diseases and disorders			
PT	Disclosure; Page 339-343; 387pp: English.			
PS	This sequence represents a novel STE20-related protein kinase. The			
XX	invention relates to nucleic acid molecule encoding a kinase polypeptide			
CC	selected from STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3,			
CC	ZC4, KHS2, SNUJ1, SNUJ3, GEX2, PAK4 and PAK5. The proteins are used to			
CC	identify agonists and antagonists, and to raise antibodies. The			
CC	polynucleotides are useful in gene therapy protocols. The polynucleotides,			
CC	polypeptides, antibodies, antagonists and agonists may be used to treat			
CC	diseases such as immune-related disorders and diseases (e.g. rheumatoid			
CC	arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.			
CC	Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,			
CC	rhlntils, autoimmunity, and organ transplantation, chronic inflammatory			
CC	pelvic disease, multiple sclerosis, organ transplantation, myocardial			
CC	infarction, cardiovascular disease, stroke, renal failure, oxidative			
CC	stress-related neurodegenerative disorders (e.g. amyotrophic lateral			
CC	sclerosis, Parkinson's disease and leigh syndrome), cancer.			

PT	Novel kinase-related polypeptides used for the diagnosis and treatment of kinase-related diseases and disorders
PT	
...	

YX

Claim 11; Page 274-278; 387pp; English.

XX

This sequence represents a novel SREB20-related protein kinase. The invention relates to nucleic acid molecule encoding a kinase polypeptide selected from S1LK, S1LK3, S1LK4, S1LK5, S1LK6, S1UK1, ZC1, ZC2, ZC3, ZC4, KHS2, S1UK1, S1UK3, GSK3, PAK and PAK5. The proteins are used to identify agonists and antagonists, and to raise antibodies. The polynucleotides are useful in gene therapy protocols. The polynucleotides polypeptides, antibodies, antagonists and agonists may be used to treat diseases such as immune-related disorders and diseases (e.g., rheumatoid arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g., Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis, rheinitis, autoimmunity, and organ transplantation, chronic inflammatory pelvic disease, multiple sclerosis, organ transplantation, myocardial infarction, cardiovascular disease, stroke, renal failure, oxidative stress-related neurodegenerative disorders (e.g., amyotrophic lateral sclerosis, Parkinson's disease and Leigh syndrome), cancer, cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes mellitus, fibrotic and mesangial disorders. The proteins may also be useful for cell growth regulation (e.g., in wound healing), T cell activation, mitosis control, and as immunosuppressants.

3
y

sequence 129 / AA;

Query Match 25.78; Score 98.5; DB 20; Length 1297;

Best Local Similarity 31.28; Pred. No. 0.041;

Matches 24; Conservative 20; Mismatches 22; Indels 11; Gaps 1.

QY 1 NRQRP---QQYECCQERCQH-----ETPRIMQTCQRCERRYEKEKRKQQKRYEE 51

[illegible]

Db 327 nkersealrrqleqqreneehkrrqlaerqkrieqkegr--rrleeqqrrekelrrkq 384

QY 52 QOREDEKEYEERMKEED 68

Db 385 qereqrhyeeqmree 401

Search completed: March 1, 2001, 15:47:09

Job time: 234 sec

